DOCUMENT RESUME

ED 023 109 CG 002 763

By-Noblitt, Gerald; And Others

Setting Up and Operating a Common Supplementary Measurement and Evaluation Service for Twenty-Three Separate School Units.

American Personnel and Guidance Association, Washington, D.C.

Pub Date 8 Apr 68

Note-19p.; Speech presented at the American Personnel and Guidance Association Convention, Detroit, Michigan, April 7-11, 1968.

EDRS Price MF-$0.25 HC-$1.05

Descriptors-County School Systems, Measurement, *Program Descriptions, Program Design, Program Development, *School Services, *Supplementary Educational Centers, *Testing Programs

This report describes a standardized testing service for multiple school districts. The testing center of the Wabash Valley Supplementary Educational Center was established through Title III of the Elementary Secondary Education Act (ESEA). The center offers four services to an eight county area testing materials, scoring, materials development, and educational services. The development and introduction of the testing program in the local schools is discussed. Changes within the local districts due to the standardized program are also presented (NS)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

AMERICAN PERSONNEL AND GUIDANCE ASSOCIATION CONVENTION

Cobal Hall Room 3045
Detroit, Michigan
April 8, 1968

Program: Setting Up and Operating a Common Supplementary Measurement and Evaluation Service for Twenty-Three Separate School Units

Chairman: Gerald Noblitt, Director
Measurement and Evaluation Division
Wabash Valley Supplementary Educational Center
Indiana State University
118½ North 6th Street
Terre Haute, Indiana 47801

Presenters: Gerald Noblitt
Topic: A Multiple School District Standardized Testing Service

Wilburn Rowe
Measurement and Evaluation Specialist
Worthington-Jefferson School Corporation
Worthington, Indiana 47441
Topic: The Past, Present and Future of Testing at Worthington-Jefferson School

Fred Brooks
Measurement and Evaluation Specialist
Greencastle Community School Corporation
500 Linwood Drive
Greencastle, Indiana 46135
Topic: The Past, Present and Future of Testing at Greencastle Community School Corporation

Dr. Lawrence Beymer, Consultant
Associate Professor of Education and Psychology
Indiana State University
Terre Haute, Indiana 47809
Topic: Implications of this Type Measurement and Evaluation Project

ED023109

CG 002 763

A MULTIPLE SCHOOL DISTRICT STANDARDIZED TESTING SERVICE

by Gerald Noblitt, Director
Measurement and Evaluation Division

Introduction:

The Measurement and Evaluation Division of the Wabash Valley Supplementary Educational Center was developed with money obtained from the federal government through Title III of the Elementary and Secondary Education Act. Title III funds are to be used to set up and operate exemplary and innovative programs which will benefit the educational program of local school corporations and might be of the type that the local school corporations would desire to continue to finance after the Title III funds are run out. Right at this point, most of the money that is spent in this project comes from the federal government through the North Vermillion School District and it is channeled to our project through Indiana State University.

There are several other divisions in the Wabash Valley Supplementary Educational Center besides the Measurement and Evaluation Division. These Divisions are: Executive Division, Fine Arts Division, Indiana History Division, Language Arts Division, Music Division, Program Development Division, Programmed Instruction Division, Pupil Personnel Services Division, and Team Teaching Division. Some of these divisions operate demonstration centers; others are more service oriented, such as the Measurement and Evaluation Division.

We serve an eight-county area here in western Indiana, as is shown here on the screen. This eight-county area has twenty-three school corporations, with a total enrollment of approximately 55,000 students. Vigo County here represents approximately one-half of our population.

With the exception of Vigo County School Corporation, all of the corporations are under 5000 students in size, with most of them in the range of 1000 to 2000 students. There are some economic and social limitations in this eight-county area which have combined to create a real need for a service of this type.

Here in Greene County, we have seven small school corporations, ranging in sizes in grades 1-12 of about 450 up to the largest, which is about 1500 students. In many of these corporations, the methods used to evaluate the educational achievement of students could be improved.

The overall objective of the Measurement and Evaluation Division is to improve the methods used by all, including students, to evaluate the educational obtainment of students. This includes the use of standardized group achievement and academic aptitude tests,

teacher developed tests, and data collection instruments and questionnaires of various sorts. There is much stress on the correct use and interpretation of these instruments.

We are trying to develop within each corporation, one person who would be confident and able to help fellow staff members, including both teachers and administrators, develop better methods for evaluating the educational obtainment of students for use within their schools and classrooms. This person would conduct workshops and inservice meetings dealing with topics from operational objectives of classroom instruction to the relationship between student achievement and social-economic level. He would operate the school corporation's standard testing program and would be looked to by the teacher staff and administrative group for leadership in all areas of evaluation of pupil academic obtainment.

Four Services

The Measurement and Evaluation Division is offering four services to local school corporations. They are testing materials for loan or sale, scoring services, materials development, and educational services. No school corporation is forced to use any of these services; however, any one or all four of these are available to each of the school corporations within our area.

We loan and sell test booklets, manuals, answer sheets, keys, and other commercially prepared testing materials. We have two academic aptitude tests and three achievement batteries for loan, free of cost to our schools. These tests are the Lorge-Thorndike Intelligence Test, Differential Aptitude Test, Iowa Tests of Basic Skills, Tests of Academic Progress, and the Stanford Achievement Batteries. These tests were selected from the better commercially prepared group tests by the Measurement and Evaluation Division Advisory Board, the division director, and two or our consultants, Dr. Lawrence Beymer and Dr. William Osmon of the Indiana State University faculty, who were involved with writing up the project.

During the planning grant stages of this division back in 1966, many of these school corporations were found to be using either outdated tests, poorly prepared tests, or poorly normed tests. The oldest test that was found in use was an edition of the Essential High School Content Battery that was standardized back in 1938.

Our scoring services include both the regular commercial MRC service and an IBM 1230 scoring machine that is located at Indiana State University. The bulk of the job is being done by MRC in Iowa City. From them we are buying a good complete service. The IBM 1230 is used mostly with scoring locally prepared tests and the Stanford Achievement Batteries.

In order to encourage local school corporations to adopt a good testing program from our testing materials, we only made the MRC scoring services available with certain tests at certain grade levels and at certain times of the year. The MRC service is available through us only with what we have come to term our Basic Supplementary Service which is shown here on the screen. Most of the schools are using this service. We are also developing local norms for each of these tests.

Our Alternate Supplementary Service uses the IBM 1230 scoring machine. The Alternate Supplementary Service is used with the Stanford Achievement Test Batteries. In the future, we hope to get more school corporations to use the IBM 1230 and our 360 computer to help them develop some semi-professional achievement tests, which are developed by them from their own educational objectives. The 1230 will also be used with other data collection the local school corporations might be interested in developing.

Whether we score the answer sheets at ISU or MRC, we get an IBM card punched with the test scores of each child on each test. This includes their raw scores, as well as interpreted scores. In this manner we are establishing a local data bank on each of these schools. We will use this bank to develop such things as local norms, and comparative studies of different types that would be of interest to local school corporation personnel.

Materials development goes right along with our data bank and comparative studies. In the future we expect to develop some expectancy tables. An example that we felt is particularly promising, is the use of the 10th grade Lorge-Thorndike Intelligence Tests scores and/or Academic Aptitude Tests scores, plus perhaps the grades earned by the student during his first semester in high school, to develop an expectancy table to help predict a student's first semester performance at ISU. Many of these students will come to ISU.

The most promising development that we hope to have out this fall is our pupil profile chart, shown here on the screen. It will be pre-printed with the scores coming from our data bank. Please note the standard errors of measurement that are being printed. Vigo County School Corporation, that has a total enrollment of approximately 23,000 students, is extremely interested in this. They have had many problems in trying to communicate the Iowa results to parents. We are planning on having these profile charts out for the Iowa, Tests of Academic Progress, Differential Aptitude Tests and the Lorge-Thorndike Intelligence Tests.

The next item that we are developing for our achievement tests is what I have come to call a Usable Item Analysis Report. Please look at our sample copy. These are available from us by writing to me. It will be used with our MRC item anlaysis reports to help local school people have a better look at what our commercially prepared achievement tests measure and how their students do on the different areas within the tests. These reports are available for use at the classroom level, building level, and corporation level.

There are many other materials that we are developing that are worthy of note. Enlarged answer sheets printed by a poster board company are planned. A picture of a smiling gear for use with primary children is planned. Research reports, dealing with studies that we do with commercially prepared tests, are planned. Tests reviews are planned. Locally developed data collection instruments are planned for use, as well as commercially prepared ones like Houghton-Mifflin's, Harcourt, Brace, and World's, and S.R.A.'s. A machine scorable high school follow-up instrument is being planned for use with our 1230.

There is a great need for what I have called our educational services. I have grouped the educational services under four topics. First are workshops that are attended by the local Measurement and Evaluation Specialists from each corporation. These are conducted by myself and the ISU Measurement faculty. This will also include this coming year, local corporation workshops, which will be conducted by each Measurement and Evaluation Specialist, using myself as a resource person. These will be held either during the week with Measurement and Evaluation Division hiring substitute teachers where needed, or on a Saturday with an honorarium being paid to most of those in attendance. Most of you that are familiar with Measurement and Evaluation, I believe, will recognize the big job that lies ahead in this area.

Inservice training is to be carried out by local Measurement and Evaluation Specialists. They are dealing with such problems as test administration, answer sheet organization and clean-up, and test scoring interpretation. Teacher-made tests and operational objectives should also be included.

Consultant services are available to any of our local school corporations. These should be used to consult on such topics as overall testing program, our services, operational academic objectives that are measured by our achievement tests, etc. These services are used to help support the Measurement and Evaluation Specialists in his role and to help the corporation itself.

An information service is also used. This includes locally written and printed monthly newsletters, bulletins, commercial reports and test bulletins, etc.

Special Problems (as I see them)

These are the big problems that we have had as I see it and I want to stress that it may only be as I see it. First was the lack of a well-organized, prepared kick-off of the Measurement and Evaluation Division. Meetings were held with local school personnel, then a lapse of time, more meetings and more lapse.

The tests and scoring services that were going to be available kept changing. Some local Measurement and Evaluation Specialists really wondered if any thing was ever going to really be done. Or if it was done, would it be done correctly. I would say the biggest problem was just plain the lack of a director for the project, which related very directly to the salary that was being offered for the position considering the longevity of the job. The project is under local school board control and the salary being offered was small to attract an experienced person. So this was the big problem which created many other problems.

The funding date of the project was during the winter of 1966-7. Well, no one was available at that time for employment.

Anyway, it was quite a job trying to help each corporation and trying to keep the promises that had already been made.

The next problem was directly related to the first problem discussed. And it was the lack of understanding of local school personnel, including Measurement and Evaluation Specialists, Superintendents, and others about just what the role of a Measurement and Evaluation Specialist within that corporation might be or what it should be to meet the needs of their corporation. The spade work to inform the superintendents about what a well-informed Measurement and Evaluation Specialist could do for their corporation just wasn't done. And this is one thing that we will be stressing this coming summer at our M & E workshops. In some cases, the person who servided this past year will be leaving the corporation. We have had problems getting materials returned so they can be mailed to others. Answer sheets in some cases have been in very poor condition, including some that looked like they were rained upon. Answer sheets have been filled out in ink or not completely filled out. Birth dates have not been included, filled out or recorded incorrectly.

The Measurement and Evaluation Division has also faced some administration problem. One was how to sell answer sheets and other testing materials to local school corporations without excepting any money for them. The ISU bookstore came to our rescue and we are using their money and billing services. This has been a great deal of help in our application for re-funding because it helped to illustrate how the local school corporations are contributing to our program themselves. It also saves a lot of answer sheets at 7.2¢ per sheet and up from being wasted.

Another problem that had to be faced was the lack of travel money. I have made two over night out of state trips to attend meetings and consult with others and have had to pay for them out of my own pocket because of the lack of travel money. The ideas I picked up certainly helped the division but it certainly didn't help my pocketbook.

I want to end on an optimistic note though. After being with the project only nine months, I believe I have learned a lot. And with a few changes here and there, and a new budget to work with, I feel our services will be much better this coming year.

Please feel free to write or come down and visit us. My address is on the last page of this booklet.

THE PAST, PRESENT, AND FUTURE OF TESTING AT WORTHINGTON-JEFFERSON SCHOOL

Wilburn H. Rowe

BACKGROUND

To introduce myself, I am Wilburn H. Rowe, an M & E Specialist, representing one of the twenty-three Wabash Valley Supplementary Educational Center school corporations, namely Worthington-Jefferson School Corporation, which is located some sixty miles southwest of Indianapolis, Indiana in Greene County. We, the Worthington-Jefferson School Corporation, are the smallest of the twenty-three school corporations making up the testing Center, with a total student enrollment of approximately 450 students in grades one through twelve.

Here is a background as to why we needed such an organization as the WVSEC testing service. Greene County itself is classified as a depressed area and is one of the few counties in the state with a declining population. Our particular area is classified as agricultural, yet the majority of our population works in industries in surrounding cities. We are a town of only 1700 population.

In Greene County we have seven separate school corporations, with seven distinct school boards. The enrollment in grades one through twelve for the total county is less than 6000 students. The school corporations in Greene County range in size from ours with approximately 450 students to about 1500 students. Our per student capital wealth in Greene County ranges from a low of approximately $5000 of assessed evaluation behind each child in the most depressed corporation, to a high of about $12,000 in the more affluent corporations. As children reach adulthood, they leave Greene County to look for work. As a result, we are heavily loaded with people in the older age bracket.

OBJECTIVE - Provide services which the small schools could not provide for themselves

This socio-economic picture of our area in Greene County, makes it quite obvious why we needed a Center such as the Wabash Valley Supplementary Educational Center, which could coordinate all the testing from the small schools within our county, as well as the schools in the other seven surrounding counties.

With the Center organized as it now is, with a student enrollment of some 50,000, we are now able to do what many larger corporations have been doing in the areas of testing.

Prior to the induction of the Center we were doing some standardized group testing but what testing we did was largely from tests that were obsolete. We had a poorly planned testing program and now we realize we were over-testing in some areas and under-testing in others. After much of the testing was done, little was done with scores except filing them away. The tests that we used were mostly those that we already had and because of the lack of money, we worked with what we had.

However, with the advent of the Measurement and Evaluation Division and with its resources (testing materials, scoring services, special materials, educational opportunities, etc.) available to us, we have been able to study, evaluate, and reorganize our testing program. There were many difficulties to be resolved in starting a testing service so that the Center would be of maximum value to the twenty-three school corporations within this eight-county area. Some of these differences that had to be taken into consideration were:

1) The testing needs of the schools in the various corporations were different.

2) The attitudes of the numerous superintendents, principals, testing specialists, and faculty toward testing varied.

3) The testing histories of these corporations were different.

4) The role of the Wabash Valley Supplementary Educational Center would be viewed differently by different superintendents, principals, and testing specialists, and even school boards.

5) The roles of the local M & E Specialists would be interpreted differently by each corporation.

6) The philosophy of the various schools in regard to testing differed.

7) The various concepts of the use and distribution of test results would be different from corporation to corporation.

8) Some corporations already had committments or had purchased up to date testing booklets and materials but different than those recommended by the test adoption committee.

9) A wide variation in the educational experience of the M & E Specialist for working with tests and test results.

All these differences had to be absorbed or at least considered in setting up a useful and workable program.

The Measurement and Evaluation Division set about through the use of the University consultants, such as Dr. Beymer, and its Advisory Board to bring together the better ideas on testing and tests. These people worked together to create a set of needed and usable testing services.

INTRODUCTION OF PROGRAM LOCALLY

Now, to the introduction of the program into the Worthington-Jefferson School system. Locally the program had a poor introduction. The program was talked about, then a lapse of time, another bit of information, another lapse and so on for about a year. Due to this delay we were not in very good shape for a well-organized kick off and inservice training program when the time came. We suffered some setbacks that could have been avoided had we waited until a given time and presented a well-organized program all in one package. The way that it has turned out is that the program has just grown with its concepts and misconcepts.

Starting out with this poor introduction has meant that the selling of the program has been more difficult than might have been could it have been inaugurated in one package deal.

PROGRAM ADOPTION

Since change is usually terrifying and since we had some changes to make in our testing program, we had and have a few problems:

1) We changed from Spring achievement testing to Fall testing. All teachers, expecially elementary, feel proud in the achievement of their students, even the lowest achievers. They felt the growth could not be detected (or their effort could not be appreciated) if the achievement testing were done in the fall. However, by pointing out that a greater insight could be gained into the needs of the children for the ensuing year, this obstacle was overcome.

2) Instead of yearly testing we went to alternate years. This again caused some concern and I think even a let down, because again the teachers did not feel that the children's growth (or the teachers' effort) could be measured.

3) One of our selling tasks was in the lower elementary grades. In order to take advantage of the scoring service offered by the Center, we had to use the MRC answer sheets. I am sure all of you are familiar with the MRC answer sheet and the more or less tedious method of marking by darkening in the small ovals. For a third grader who has been accustomed to using the wide lined paper and making big letters and his coordination being what it is, the third grade teachers were rather reluctant to give their consent to change from consummable booklets to the MRC answer sheets for machine scoring. This problem of selling the third grade teachers was overcome by a compromise deal: let's try it once and see. This along with the idea that they (the teachers) didn't have to spend hours at home grading the tests led to their consent. Now that we have tried the MRC answer sheets and the children did well, I believe the teachers are now sold on the MRC answer sheets. We do feel, though, that a more usable MRC answer sheet should be devised for use in grades three and four. If this is not possible, then the MRC practice sheets should be considered prior to the first use of such machine scorable answer sheets.

ADVANTAGES OF TESTING CENTER

We do feel that the testing Center has many advantages, namely:

1) This eight county center with its 50,000 student enrollment has made us a large corporation, test wise, with all its advantages for securing tests, scoring services, developing materials, and holding well-organized workshops.

2) Economy of the program - our testing now costs us approximately 7.2¢ per answer sheet, in contrast to about 20¢ to 25¢ plus our booklets, before the advent of the Center.

3) Our staff is better informed about testing materials and test interpretation than before.

4) The people in the testing Center, such as Mr. Noblitt and the consultants with whom he works, are specialists in their fields and can keep more up to date than we in the local schools who have a variety of non-testing related duties to perform. With these contacts and our periodic workshops we too can be better informed and through this have an improved testing program.

5) Through the workshops which are held by the Center for all twenty-three Specialists, we have time to discuss difficulties we are having that are of common interest. There are always an assortment of ideas on how these difficulties can and should be approached or solved and through the workshops we learn these ideas.

6) We feel that the consultant and inservice training programs are a great asset.

7) Ease, speed, and economy in procuring testing materials has improved greatly as a result of the Center.

8) Through the basic and alternate services, the eight-county schools have enough uniformity for comparison and norming and yet enough variation to tailor the testing program to the individual school choices and differences.

9) More uniformity of test programs will help us do a better job of helping those students moving from place to place within the eight-county area.

10) We have the opportunity to hear from some of the people who do our scoring and from them find out about some of our errors in administering the standardized tests.

CONCLUSION AND PROBLEMS

We have many problems remaining, many of which we probably don't even realize as yet, but many we do know. One of our major problems, at least in our local school, is that of score reporting. We have such questions as:

1) To whom should scores be revealed or reported?

2) If scores are reported, to whom and how should we report them?

3) If scores are or are not reported, then what use are we to make of them?

Another problem facing us is how best to use the available consultant services and what and how to use the inservice training programs.in order to economize on time, money, and gain the most from such services and programs.

Fred L. Brooks

The Past, Present and Future of Testing at Greencastle Community School Corporation

Introduction:

I am Fred L. Brooks, Licensed Psychometrist and Director of Testing for the Greencastle Community Schools and Psychometrist (School Psychologist) for the Joint Special Services Program in Putnam County, Indiana.

I. Setting

The Greencastle Community School Corporation includes Greencastle and Madison Townships in Putnam County serving over 2500 students in four elementary schools, one junior high, and one senior high school.

Greencastle is a county seat city of 10,000 population with the cultural advantages of DePauw University, a liberal arts college founded by the Methodist Church in 1837. Five major industries and several minor industries together with a score or more thriving businesses surrounded by successful farm operations in the rural areas make Greencastle and Putnam County high in wealth and resources compared to cities and/or counties of the same size and population in Indiana. In spite or this there are two poverty areas in the city of Greencastle, on the south and east, where the cultural lag necessitates headstart programs to bridge the gap. Statistically there is a good balance of the extremes as well as a substantial representation of the average.

II. The testing program in the Greencastle Schools consists of readiness at the end of kindergarten, achievement testing in grades one through six in the elementary, grade seven in the junior high and grade ten in the high school. Group intelligence tests are given in grades three, seven, and ten. An aptitude test is given high school freshmen and an interest inventory is administered at the beginning of grade twelve. The usual college entrance and scholarship tests are given to juniors and seniors. Individual testing is available on referral for intelligence, personality, and special problems such as screening for brain damage or emotional disturbance.

III. Before the Wabash Valley Supplementary Educational Center came into being with one of the divisions devoted to Measurement and Evaluation, there were many tests given in the Greencastle School Corporation. In some instances they were not even scored and recorded.

Since I joined the staff in September 1962, the tests have been scored and recorded, but it was discovered that little use was made of the results. The greatest contribution made by the Supplementary Educational Center to my corporation has been in

the area of interpretation. Workshops involving the Measurement and Evaluation Specialists from each of the twenty-three school corporations have emphasized the purpose of testing, the proper selection of a testing program to achieve various purposes (i.e. achievement of pupils, placement, curriculum evaluation), and the interpretation of test results in statistics and form practical for teachers, counselors, administrators and parents. It was in this area of interpretation that I feel the greatest contribution has been made for our corporation. The basic program offered by the Center did not supplant our present program. It did supplement as the term of the Center implies. It did enrich the program. Out of it came a confidential statistical report to the superintendent, principals, and counselors; screening for remedial programs and bulletins to teachers on the interpretation of scores reported on past labels in the individual records. New possibilities in the use of scores were revealed. In the interplay of ideas in the workshops each participant was made more aware of problems and approaches to solutions regarding the whole gamut of the testing process. Even testing companies were given instructions as well as more complete reporting in grade scores, stanines and percentiles so that each teacher could select the score most meaningful to him and note its comparison with other scores.

The most recent development in the Center has been the introduction of item analysis to be used by individual teachers for instructional purposes in areas where individuals or groups of students are weak; or use by the administration for curriculum evaluation and study.

IV. Future Plans

It may appear that everything has been done that needs to be done. But this is just the beginning. Next year the testing program will be planned by a committee after a study has been made to determine what purpose the tests are to serve. There will probably be less testing and more emphasis on their uses. These suggestions came out of the Center. There will be inservice training for teachers and counselors so they will be able to suggest the type of tests needed to fulfill their purposes and to be more competent in score interpretation for their own uses or in sharing results with parents. These suggestions came out of the Center.

With such improvement of the testing program in the Greencastle School Corporation, the needs of the pupils can be better identified and steps taken to improve the instructional program as well as screening special needs to better meet the individual differences of the children placed in our care.

If you're interested in keeping up with further developments in our testing program, I would like to correspond with you.

IMPLICATIONS OF THIS TYPE
MEASUREMENT AND EVALUATION PROJECT
by Dr. Lawrence Beymer

I would like to call your attention first to the fact that it took the United States forty years to work its way through the industrial revolution; Japan did it in eight. The point of this is that there really isn't much use for you people to make all the stupid mistakes that we have made; we've already done it. I believe this is entirely within the spirit of the project and if you are considering this, then I think one of the best things you can do is come down and visit us for a couple of days. We can show you things that you wouldn't believe.

We hope it is not a product of too much stupidity but simply somewhat like what Theodore Greene, a senator from Rhode Island who died at the age of 96 or 97 once said about politics, "Politics are a lot like roller skating; part of the time you go where you go because that's where you want to go and part of the time you go where you go because that's where the darn things take you." So we just kind of got on and the things just kind of go.

We had trouble getting started. The federal government in all its wisdom just doesn't realize that the school year starts in September. By the time that they said go, many schools had started their own testing programs. People whom we might have been able to contact for leadership and operational managerial skills were not yet available or couldn't wait for us and took other jobs. Therefore, we really haven't been in business as long as it might look like on the official record. To tell the truth we haven't probably been really in any more than low gear for more than about one year.

I think that what you have heard today demonstrates, I hope, that we never started to develop a program which would supplant all the schools' efforts in measurement in these eight counties with between 50 to 55 thousand boys and girls. We couldn't possibly do that. We could certainly supplement and that's what we're trying to do. We're trying to bring some order to it, some predictability to it, some structure to it, with enough freedom for elaboration so local adjustments could be made.

We've had some interesting problems. One is that from the very beginning when I began to write the proposal, we made the statement that we're no longer going to get by with the old idea that we can't do anything because there's nobody out there who is trained. We said that there are people out there who are interested, who have the ability, but who need to be updated. And we began to do this the very first summer. We brought these people together, not as inferior, but as colleagues facing a common task to see if we couldn't get together, refresh ourselves, sharpen ourselves up a little bit, and get ready to go to work.

We've discovered some interesting things about measurement and guidance. It became very clear to us that for some reason it's the school counselor who's expected to be the measurement specialist. This fascinates me in a way. I think our institution also takes the point of view that it's really somewhat poor taste and bad sportsmanship to complain all the time about how people aren't really qualified to do things, when they have degrees from your own institution. The fact of the matter

is that in typical schools we know about in our section of the country, the counselor probably has as much training as any two people put together. And I don't understand why this is true. The typical school administrator as we know him doesn't even suspect too much about measurement. It hasn't been in his background. Many professional school people held beliefs about tests that just weren't true. We found out that we couldn't take anything for granted.

I think counselors are involved in measurement because measurement is a real paradox. It looks so scientific, it looks so quantitative (after all, we use decimal points), that it's one of the most subjective and emotional things that you can ever start tinkering with. When you quantify something, you are making indirect implications about an individual's entire environment -- his parents, his school, the quality of teaching, and the effort his local community has put forth in the school. For instance, it's a shock to some of our people to understand that when they go to local norms, half their people have to fall below the middle.

If you select a well-normed standardized test with care, the typical middle class school can't punch 75-80% of their students above their median. But this is a shock and we had to react to this shock. A school may be quite proud of its math department. It finds out when we give the Academic Promise Test that half of their kids are below the middle. They understand that intellectually but not emotionally and so we have to help deal with this.

You have the beginnings of this business of item analysis. What happened? Right off the bat the people said, "My God, only 44% of our students answered this first item correctly. Twice that is barely satisfactory!" Sometime in the past, I'm sure someone told them that when you build standardized tests, you pitch out all the items that everybody gets right, and you pitch out all the items that nobody gets right, and most of the items left are those that 40-60% of the norm will get right. And they heard it and they wrote it down but that shows you how something can go from your ears to your arms, from your pencil to your paper, and not go to your brain. They just don't remember that. They get quite upset and we have to help them understand.

We had some problems with test publishers and salesmen. I had a salesman who came to see me Friday afternoon - quite upset. He says that we're destroying his business in western Indiana. I answered him that I don't care any more about that than I care about lots of other things. We do not own stock in Houghton Mifflin. We're not really very happy with all the things they do either. I'm not very happy about this grade placement gimmick that they use, which is a grade placement with no decimal point, with the assumption that this is interval data and the distance between 70-80 is the same as with 30-40, and that's not true. You go downstairs and ask Ed Drahozal and he'll tell you that it's not true either.

I think that measurement and professional people face a dilemma that's going to be quite painful. We have a head-on clash between salesmanship and scholarship and it's a mis-match. We've had test salesmen come to us trying to get us to buy and use their tests, who have told us that the typical school teacher is too dumb to understand anything but grade placement, therefore that's what we ought to give them. I don't believe that for a minute. It's not true.

We've had a little trouble. But we intend to expand our services; once we get the bugs ironed out and the computer printing out standard errors of measurement and percentile bands on profiles. It may be that we will add another high school achievement test and another elementary school achievement test. Our main thrust here is toward measuring achievement, measuring status, and doing diagnostic work. We hope sometime in the future to have a situation where these men can sit down with the teacher and they can say, "You have five kids in your class who made this score but they're not alike." Billy made 3/4 of his errors on the math test on the items that had anything to do with fractions, and he didn't understand them very well. I looked back in his history and found out that Billy had rheumatic fever once in the fifth grade and was out of school for seven weeks. He just never learned fractions. So maybe you can help a little here and give some practices exercises. I think the purpose of all this is to emphasize the uniqueness of each of these children so that they can be taught differently than they would have been had you not known this information; hoping that the next time they come around to test, we'll be able to look at Billy's math score and say, "Well, he's still not getting them all right but no one else is either. But at least he's getting more of them right. Billy as compared with Billy is gaining a little ground." So these are the type of things that we hope to grow in.

We were not kidding ourselves when we began. It takes a long time. These men up here have lots of other things to do besides fool around with tests and so do I. But I think that now we're beginning, at least in our section of the state, to gain confidence in what we do know and are finding some success and we're now at the place where we are beginning to move and make a difference. And to that extent we have been successful.

We are constantly engaged in evaluation of this project. Every couple of months, we sit down and find out what went wrong this month, what problems arose, what we should do that we can't do, what were alternatives, which one did we choose and then what happened? That stuff is all printed down. So again, you need not, if you are interested in setting up this type of project, make our errors. Make some of your own.

QUESTIONS and GENERAL DISCUSSION following Presentations

QUESTION: We're about to begin a rather large scale testing program and we're doing our planning now. One of the big questions that we have delayed judgment on is "how to report the scores!" We're very strongly moving in the direction of reporting nothing but raw scores, and in addition to that, developing our own local norms and making these norms available. Now in your program, I noted that you are providing percentiles. We have a problem knowing which reference to base percentiles on. Of course, that reverts you way back to the raw scores and the assumption that if the people can't use the raw scores, they shouldn't be using the tests at all. Do you have any guidelines for us?

ANSWER (Beymer): The best way to demonstrate the fact that tests are not absolute indicators is to always report scores either in two statistical derived ways or a couple of norm groups. We found out that taking grade placement away from teachers who really think that all kids must be 7.0 to be in the 7th grade gives them withdrawal symptoms. We couldn't accept that. So what we are trying to do now is to give them what they are used to, plus something new, side by side, and in a couple years that need for grade placement scores is going to fade away. I think it's a mistake not to give them what they are used to. In norms, we can give them national norms, area norms, district norms, and school norms. Next fall, Mr. Rowe can tell other teachers, "Well look, the reading ability of our 7th graders is about like this this year, and it was about like that last year." He can compare, if he wishes, how his school's results compare with the results in western Indiana among schools which are playing our game but he will not know how his students compare with the students in Greencastle. No school scores are ever identified out by the school. So we can say your freshmen are about like this and the freshmen in western Indiana about like so. So one way that we are trying to do this is to give people percentiles and stanines, or percentiles and grade placement on the Iowa as it comes back. Then maybe local norms and national norms together. Now they become confused by this. It's a confusing business. And I think that you don't do people a justice when you oversimplify something that's kind of relative. With whom would you like to compare this child, madam? The way he was two years ago or with the group that he finds himself a part of now or with the kind of kids that we sent you last year from our school. What would you like to know? And so much of this involves helping teachers learn how to ask what they want to know about. When you go to the English department and you say, "Here's how they're alike and how they're different." Then you see, all of this stuff they say about how "they would teach them better if they knew what they already learned but they don't so let's go ahead" changes color. All of a sudden now, once they know

some things, they can make an adjustment. So I would hope that we could give a couple of things and if you find yourself in a place where teachers absolutely insist on something, I think that you would be ill-advised to simply refuse to report the scores this way. But maybe two ways. They can learn them.

(Noblitt): I'd like to talk just a minute on that. As most on you know, the Lorge-Thorndike test has two different norms that they report on the gum-labels and in the manuals. One is the age norm group and the other one's the grade placement. I haven't got it developed yet, but I do hope to have a pupil profile chart out such as this one right here which will report both of them. We just label them both and the computer can print the standard error of measurement out by these scores. I think a person could do that with any type of score they would want with reference to any norm group they would want.

QUESTION: What's your total budget?

ANSWER (Noblitt): Our total budget is about $65,000 per year.

QUESTION: For how many years?

ANSWER (Noblitt): That's per year.

QUESTION: How long have you been in operation?

ANSWER (Noblitt): We've been in operation about a year.

QUESTION: Through your funding, was there any local district contribution?

ANSWER (Noblitt): In theory, but many of the corporations would have a hard time supporting their share of the cost. They are helping to support part of the project by buying their own answer sheets, but that isn't a very large amount.

(Beymer): Also by salaries of these local people on the job.

QUESTION: What is the expense per pupil?

ANSWER (Noblitt): Divide $65,000 by 53,000 pupils. It's true not all those now are being tested, but some groups have two tests per year, so I think we're averaging about $1.25 per pupil. This will probably go down. But I'm not sure that's the good way to look at it. Our workshops and such are really taking a lot of our money, getting people trained.

(Beymer): We're trying to put most of this thing into developing local people to handle it so it's sort of capital investment, we feel. A lot of school systems are putting their money into computers. Don't put your money into

computer and data processing and neglect the firing line out here -- the people who have to consume it. You have to begin to develop their ability to utilize the results. We've tried to move slow; we haven't gone around and really done a hard sell job. We would be shocked into paralysis if every school in these eight counties all came in and all wanted everything they're technically eligible to get. By next year, we'll kind of be going around and saying, "Hey, we got some of the bugs worked out." Some of them have told us to practice on Worthington and Greencastle and then let them know and maybe they'll play later.

QUESTION: Do you have any plans to develop expectancy tables within this program?

ANSWER (Noblitt): Yes we do. At the 10th grade level, we are interested in using our Tests of Academic Progress with our Lorge-Thorndike scores, and/or the grades of students earned in their schools to develop expectancy tables related to their success here at Indiana State University during their freshman year. We have established a data bank with our scoring services and their grades here at ISU will be available, as many of the students will attend school here.

QUESTION: The college board at ETS has spent a lot of effort to provide regression equations and expectancies for Indiana universities. And haven't these been enthusiastically accepted by Hoosiers?

ANSWER (Beymer); Yes, I think that is true. I think the people are enthusiastic about what they've done here in the state of Indiana. I personally don't think much of that project. It really isn't what it was cracked up to be. For instance, what good is it for you to know expectancy tables for somebody from Worthington-Jefferson High School at Purdue University? What is he majoring in, in home ec or elementary education or electrical engineering? I wonder!

QUESTIONER'S COMMENT: I thought they examined a wide variety of criteria.

(Beymer): Yes, but I think it makes a batter report than it does a utility. To put on his desk a regression equation is about as complicated as interpreting a test and saying, "Well, keep in mind, Billy, that while your stanine score is 7, your reliability coefficient is only 26." I just don't believe this is very helpful!

QUESTIONER'S COMMENT: Yes, but I would much rather know what are the chances in a wide variety of alternatives as opposed to what his IQ is.

(Beymer): OK, but I think that the number one justification for measurement is description, not prediction.

QUESTIONER'S COMMENT: I disagree.

(Beymer): Read Robert Ebel's book. (Tape ran out.)